(19)日本国特許庁（ＪＰ）　(12)公開特許公報（Ａ）　(11)特許出願公開番号
特開2000－81894
（P2000－81894A）
(43)公開日　平成12年３月21日(2000.3.21)

(51)Int.Cl.⁷	識別記号	ＦＩ		テーマコード（参考）
Ｇ１０Ｌ　15/06		Ｇ１０Ｌ　3/00	５２１Ｆ	
15/10			５２１Ｎ	
			５３１Ｇ	

審査請求　未請求　請求項の数11　ＯＬ　（全 13 頁）

(21)出願番号　特願平11－248458

(22)出願日　平成11年９月２日(1999.9.2)

(31)優先権主張番号　０９／１４８９１１

(32)優先日　平成10年９月４日(1998.9.4)

(33)優先権主張国　米国（ＵＳ）

(71)出願人　000005821
松下電器産業株式会社
大阪府門真市大字門真1006番地

(72)発明者　クーン　ローランド
アメリカ合衆国　カリフォルニア州
93110　サンタ　バーバラ，コーレ　シタ
3928

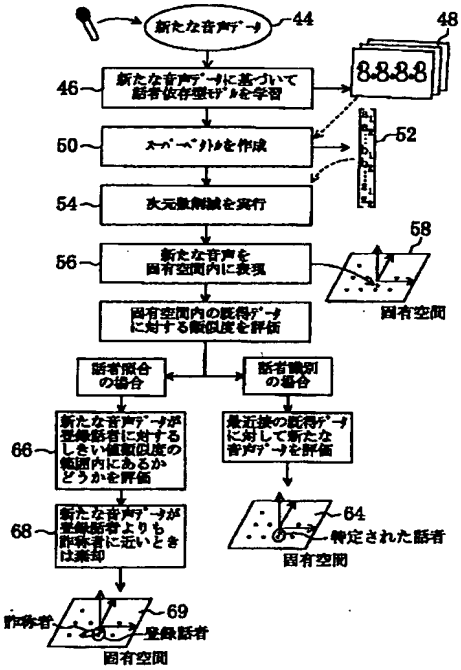
(72)発明者　ニュイエン　パトリック
アメリカ合衆国　カリフォルニア州
93117　イスラ　ヴィスタ，エル　コレジ
オ　ロード　6739

(74)代理人　100077931
弁理士　前田　弘　（外１名）

最終頁に続く

(54)【発明の名称】　音声評価方法

(57)【要約】
【課題】　話者識別および話者照合方法を提供する。
【解決手段】　登録話者の音声に基づいて（話者照合の場合にはさらに詐称者の音声にも基づいて）音声モデルを作成し学習する。音声モデルからのパラメータを連接してスーパーベクトルを定義する。スーパーベクトルに対して線形変換を施して次元数を削減し、低次元空間（固有空間と呼ぶ）を生成する。学習用話者は点または分布として固有空間内に表される。その後、試験用話者からの未知音声に対して同様の線形変換を施して固有空間内に位置づける。固有空間内の試験用話者の学習用話者に対する類似度によって試験用話者を認識する。



1

【特許請求の範囲】

【請求項1】　登録話者に関する音声評価方法であって、

少なくとも一人の登録話者を含む複数の学習用話者の音声に基づいて、音声モデル集合を学習するステップと、

前記音声モデル集合の次元数を削減して基本ベクトル集合を生成し、この基本ベクトル集合により定義され、かつ、前記複数の学習用話者を表すための固有空間を構築するステップと、

前記登録話者を前記固有空間内に第1の位置として表すステップと、

新たな話者による入力データに基づいて新たな音声モデルを学習し、この新たな音声モデルの次元数を削減して前記新たな話者を前記固有空間内に第2の位置として表すことによって、前記新たな話者による入力データを処理するステップと、

前記第1の位置と前記第2の位置との類似度を評価し、その評価を前記新たな話者が前記登録話者か否かの指標として用いるステップとを備える、音声評価方法。

【請求項2】　請求項1に記載の音声評価方法において、

話者識別を行う場合には、

前記複数の学習用話者は、複数の異なる登録話者を含み、

前記音声評価方法はさらに、

前記複数の登録話者の各々を前記固有空間内に学習用話者の位置として表すステップと、

前記第2の位置と前記学習用話者の位置との類似度を評価し、この評価の少なくとも一部に基づいて前記新たな話者を前記複数の登録話者の中から選択した一人として識別するステップとを備える、音声評価方法。

【請求項3】　請求項1に記載の音声評価方法において、

話者照合を行う場合には、

前記複数の学習用話者は、固有空間内に第3の位置として表される少なくとも一人の詐称者を含む、音声評価方法。

【請求項4】　請求項3に記載の音声評価方法において、

前記音声評価方法はさらに、

前記第2の位置と前記第3の位置との類似度を評価し、その評価を前記新たな話者が前記登録話者か否かのさらなる指標として使うステップを備える、音声評価方法。

【請求項5】　請求項1に記載の音声評価方法において、

前記類似度を評価するステップは、前記第1の位置と前記第2の位置との間の距離を定めることによって行われる、音声評価方法。

【請求項6】　請求項1に記載の音声評価方法において、

2

前記学習用話者は、前記固有空間内に位置として表される、音声評価方法。

【請求項7】　請求項1に記載の音声評価方法において、

前記学習用話者は、前記固有空間内に点として表される、音声評価方法。

【請求項8】　請求項1に記載の音声評価方法において、

前記学習用話者は、前記固有空間内に分布として表される、音声評価方法。

【請求項9】　請求項1に記載の音声評価方法において、

前記新たな話者による入力データを処理するステップは、

前記入力データを用いて確率関数を生成しその確率関数を最大化することによって前記固有空間にある最尤ベクトルを決定するステップを含む、音声評価方法。

【請求項10】　請求項1に記載の音声評価方法において、

前記複数の学習用話者は、複数の登録話者と少なくとも一人の詐称者とを含む、音声評価方法。

【請求項11】　請求項1に記載の音声評価方法において、

前記音声評価方法はさらに、

前記第1の位置と前記第2の位置との類似度を周期的に評価し、新しい話者が登録話者か否かの指標としてその評価を用いることによって前記新たな話者の同一性が変化したか否かを決定するステップを含む、音声評価方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】この発明は音声処理技術に関し、さらに詳しくは、話者照合あるいは話者識別を実行するシステムおよび方法に関する。

【0002】

【従来の技術および解決しようとする課題】本人であることを認識することは、ほとんどすべての取引における中心問題である。多くの人が電話を通じて自己の預金口座にアクセスしたり自己のクレジットカードを使用したりというような秘密の金融取引を行っている。現在実施されている本人であることの認識は決して簡単ではない。政党間では、社会保障番号、母親の旧姓などの秘密情報の交換が行われているであろう。このような情報は盗まれる可能性があり、その場合には誤った認識がされることになる。

【0003】この発明の1つの局面においては、話者照合を実行するシステムおよび方法を提供することにより上述の問題に焦点をあてる。話者照合では、与えられた音声が特定の話者(ここでは、登録話者という)に属するか詐称者(登録話者以外のだれか)に属するかを決定

3

することが必要とされる。

【０００４】話者識別における問題点は話者照合における問題点と多少関係している。話者識別では、与えられた音声を既得の音声集合の中の1つにマッチングさせる必要がある。話者照合と同様、話者識別にも多くの興味深い応用例がある。例えば、話者識別システムは、音声サンプルを利用することができる話者群に関して話者による音声メールを区分することに使用されるかもしれない。このような能力によれば、音声メールシステムにメッセージを残した通話者の身元をコンピュータ画面に表示させるコンピュータテレフォニーシステムが可能となる。

【０００５】話者照合および話者識別の応用例はほとんど無数に存在するが、話者照合および話者識別の実行の解明はこれまでわかりにくいものであった。人間の音声を認識すること、特にその話者を他の話者から識別することは複雑な問題である。人は、たった一つの単語でさえ全く同じように二度話すことはめったにない。これは、人間の発声法が原因となっている。

【０００６】人間の音声は以下のようにして生成される。肺から押し出された空気が声帯を通り抜け、声門により調音され音声波が生成される。音声波は口腔および鼻腔で共鳴し、その後、舌、顎、歯、唇により言語音が作られる。これらの音声生成機構の相互作用に対して、様々な要素が影響を及ぼす。例えば、かぜにより声帯の音質が大きく変化するのと同様に鼻腔の共鳴も大きく変化する。

【０００７】人間の音声生成における複雑さおよび変わりやすさのため、話者照合および話者識別は、新たな音声を既得の音声サンプルと比較することにより容易に実行できるというわけではない。詐称者を除外するために類似度のしきい値を高く設定すると、本物の話者が鼻風邪を引いている場合にはその本物の話者を棄却してしまうことがある。一方、類似度のしきい値を低く設定すると、システムは誤った照合を起こしやすくなる。

【０００８】

【課題を解決するための手段】請求項1に従った音声評価方法は、登録話者に関する音声評価方法であって、少なくとも一人の登録話者を含む複数の学習用話者の音声に基づいて、音声モデル集合を学習するステップと、音声モデル集合の次元数を削減して基本ベクトル集合を生成し、この基本ベクトル集合により定義され、かつ、複数の学習用話者を表すための固有空間を構築するステップと、登録話者を固有空間内に第1の位置として表すステップと、新たな話者による入力データに基づいて新たな音声モデルを学習し、この新たな音声モデルの次元数を削減して新たな話者を固有空間内に第2の位置として表すことによって、新たな話者による入力データを処理するステップと、第1の位置と第2の位置との類似度を評価し、その評価を新たな話者が登録話者か否かの指標として用いるステップとを備える。

4

【０００９】請求項2に従った音声評価方法では、話者識別を行う場合には、複数の学習用話者は、複数の異なる登録話者を含む。上記音声評価方法はさらに、複数の登録話者の各々を固有空間内に学習用話者の位置として表すステップと、第2の位置と学習用話者の位置との類似度を評価し、この評価の少なくとも一部に基づいて新たな話者を複数の登録話者の中から選択した一人として識別するステップとを備える。

【００１０】請求項3に従った音声評価方法では、話者照合を行う場合には、複数の学習用話者は、固有空間内に第3の位置として表される少なくとも一人の詐称者を含む。

【００１１】請求項4に従った音声評価方法は、第2の位置と第3の位置との類似度を評価し、その評価を新たな話者が登録話者か否かのさらなる指標として使うステップを備える。

【００１２】請求項5に従った音声評価方法では、類似度を評価するステップは、第1の位置と第2の位置との間の距離を定めることによって行われる。

【００１３】請求項6に従った音声評価方法では、学習用話者は、固有空間内に位置として表される。

【００１４】請求項7に従った音声評価方法では、学習用話者は、固有空間内に点として表される。

【００１５】請求項8に従った音声評価方法では、学習用話者は、固有空間内に分布として表される。

【００１６】請求項9に従った音声評価方法では、新たな話者による入力データを処理するステップは、入力データを用いて確率関数を生成しその確率関数を最大化することによって固有空間にある最尤ベクトルを決定するステップを含む。

【００１７】請求項10に従った音声評価方法では、複数の学習用話者は、複数の登録話者と少なくとも一人の詐称者とを含む。

【００１８】請求項11に従った音声評価方法は、第1の位置と第2の位置との類似度を周期的に評価し、新しい話者が登録話者か否かの指標としてその評価を用いることによって新たな話者の同一性が変化したか否かを決定するステップを含む。

【００１９】この発明は、話者照合および話者識別のためのモデルに基づいた分析方法を使用する。モデルは、既知の登録話者の音声に基づいて作成されて学習する（話者照合の場合には、一人またはそれ以上の詐称者の音声にも基づく）。これらの話者モデルには、例えば、隠れマルコフモデル（以下、HMMともいう。）におけるパラメータのように、一般的に多数のパラメータが使用される。これらのパラメータを直接使用するのではなく、連接させてスーパーベクトルを作成する。これらのスーパーベクトルは、話者一人あたり一個作成され、学習データの話者の母集団全体を表す。

5

【００２０】スーパーベクトルに対して線形変換を行なって次元数を削減し、低次元空間（ここでは、固有空間と呼ぶ）を生成する。この固有空間の基底ベクトルを固有音声または固有ベクトルと呼ぶ。必要があれば、固有ベクトルの成分のいくつかを捨てることにより固有空間の次元数をさらに削減することができる。

【００２１】次いで、学習用データを含む話者の各々を固有空間内の点あるいは固有空間内の確率分布として固有空間内に表す。前者（点として表すこと）は、各話者からの音声を相対的に不変のものとして取り扱う点で、やや不正確である。後者（確率分布として表すこと）は、発話ごとの各話者の音声の変化を反映する。

【００２２】各話者の学習用データが固有空間内に表されると、システムを用いて話者照合または話者識別を行うことができる。

【００２３】新たな音声データが得られるとこれを用いてスーパーベクトルを作成し、次いで次元数削減を行い固有空間に表す。新たな音声データの既得のデータに対する類似度を評価することで話者照合あるいは話者識別を実行する。話者からの新たな音声について、その固有空間内の対応する点あるいは対応する分布が登録話者の学習用データに対するしきい値類似度内であるか否かを照合する。システムは、本人であっても、その音声が固有空間内にある詐称者の音声の方に近い場合には棄却することがある。

【００２４】話者識別は類似の方法で行う。新たな音声データを固有空間内に位置付け、学習用話者のうち固有ベクトルの分布点が最も近い学習用話者と結びつける。

【００２５】固有空間内において新たな音声データと学習用データとの類似度を評価することには多くの利点がある。

【００２６】第一に、固有空間は、単に選択された数個の特徴だけでなく、各話者の全体を簡潔、低次元の方法で表す。

【００２７】また、固有空間内に含まれる次元数は、元の話者モデル空間あるいは特徴ベクトル空間内に含まれるよりも一般にかなり少ないので、固有空間内で実行される類似度の計算を非常に速くすることができる。

【００２８】また、システムにおいては、元の学習用データを作成するのに使用したすべての例、発話が新たな音声データに含まれていることは必要とされない。この発明によれば、その構成要素の一部を欠くスーパーベクトルに対して次元数の削減を行うことができる。その結果としての固有空間内の分布点は話者を明確に表す。

【００２９】

【発明の実施の形態】以下、この発明の実施の形態について図面を参照しつつ説明する。

【００３０】この発明において用いられる固有音声手法は、多くの異なった音声モデルに対して機能する。ここでは好ましい実施の形態として、今日の音声認識手法において最も一般的な隠れマルコフモデル認識系に関して説明する。しかし、この発明は、例えば音素類似性認識系のような他のタイプのモデルに基づく認識系を使用して実行することもできる。

6

【００３１】この発明による話者識別および話者照合をよりよく理解するためには、話者認識システムについて基本的な事項を理解しておくことが有用と思われる。したがって、以下、隠れマルコフモデル手法について説明する。隠れマルコフモデルは、今日のほとんどの話者認識系において話者を表すために使用されているものである。

【００３２】隠れマルコフモデルは状態図を伴うモデル化手法である。モデルに含まれている全ての知識源（句、単語、サブワード、音素など）を利用することにより、いかなる音声単位であってもモデル化することが可能である。隠れマルコフモデルは、観測可能な出力の系列を離散間隔で生成する未知の処理を表現し、出力は（予め決められた音声単位の集合に対応する）いくつかの有限個のアルファベット要素である。これらのモデルは、観測可能な出力を生成した状態の系列が未知であるので、「隠れ」と呼ばれる。

【００３３】図1に示すように、隠れマルコフモデル10は、状態（Ｓ１，Ｓ２，．．．，Ｓ５）のセットと、図1に矢印で示す各対の状態間の遷移を規定するベクトルと、確率データの集まりとによって表される。具体的には、隠れマルコフモデルは、遷移ベクトルに関連する遷移確率のセット１２と、各状態で観測された出力に関連する出力確率のセット１４とを含んでいる。このモデルは、ある状態から別の状態まで一定の離散間隔で計測される。クロックタイムには、モデルは現在の状態から遷移ベクトルが存在するどの状態へも変化してよい。図1に示すように、所定の状態からそれ自体に戻るという遷移も可能である。

【００３４】遷移確率は、モデルが計測された際にある状態から別の状態への遷移が発生する尤度を表現している。すなわち、図1に示すように、各遷移は0と1の間の確率値を伴っている。どの状態からでもその状態を離れる全ての確率の合計は1である。例として、遷移確率表１２に遷移確率値のセットを掲載する。実際の実施形態では、どの状態からでもその状態を離れる全ての確率の合計が1に等しいという制約のもとで、これらの値が学習データにより生成される。

【００３５】遷移が行われるときはいつも、モデルがアルファベットの一要素を発信すなわち出力していると判断することができる。図1に示す実施形態では、音素を基準とする音声単位が想定されている。したがって、出力確率表１４で特定されるシンボルは標準英語に見られる音素の一部に相当する。各遷移の際にアルファベットのどの要素が出力されるかは学習中に覚えた出力確率値すなわち関数によって決まる。このようにして発信され

た出力(学習データに基づく) は、観測値の系列を表し、アルファベットの各要素は出力確率を有している。

【００３６】音声をモデル化する際に共通して行われることは、離散アルファベットシンボルの系列とは対照的に、出力を連続するベクトルの系列として扱うことである。したがって、出力確率は1 個の数値の場合とは対照的に、連続する確率密度関数で表現される必要がある。このように、HMMは1 個以上のガウス分布を備えた確率密度関数に基づく場合が多い。複数のガウス関数が使用される場合、図に１６で示すように、それらは一般に複素確率分布を画定するよう加法的に混合される。

【００３７】単一ガウス関数として表現されるにせよ混合ガウス関数として表現されるにせよ、確率分布は複数のパラメータで記述される。遷移確率値１２と同様に、これら出力確率パラメータも浮動小数点数を含んでいてもよい。パラメータ１８は、学習用話者からの観測データに基づいて確率密度関数(ｐｄｆ) を表現するために一般的に使用されるパラメータを特定するものである。図1 のガウス関数１６の等式で示すように、モデル化されるべき観測ベクトルOの確率密度関数はガウス密度Nにより多重化された各混合成分の混合係数の反復合計であり、この場合、ガウス密度はケプストラム係数あるいはフィルターバンク係数の音声パラメータから算出された平均ベクトルuj 及び共分散行列Uj を含んでいる。

【００３８】隠れマルコフモデル認識系の実行の詳細は、応用例ごとに大幅に異なることがある。図1 に示す隠れマルコフモデルの一例は隠れマルコフモデルを作成する方法を単に例示したにすぎず、本発明の範囲を限定するものではない。この点について、隠れマルコフモデル化の概念に関する多くの変形例が存在する。以下の説明からより完全に理解できるように、本発明の固有音声適応化技術は各種隠れマルコフモデル変形例だけでなくパラメータを基準とする他の音声モデル化システムにも効果的であるように容易に適応させることができる。

【００３９】図2 および図3 はそれぞれ、この発明の実施の形態による話者識別、話者照合を実行するための固有空間の構築を説明するためのフローチャートである。この発明の実施の形態による話者識別、話者照合を実行するために、まず固有空間を作成する。作成する固有空間は、応用例により定まる特有の固有空間である。図2 に示すように、話者識別の場合には、登録話者集合２０を使用して学習用データ２２を提供し、この学習用データ２２に基づいて固有空間を作成する。対して、話者照合の場合には、図3 に示すように、照合の対象となる一または複数の登録話者２１ａ、さらに1 または複数の仮想の詐称者２１ｂも使用して学習用データ２２を提供する。このように学習用データ２２の源が異なるという違いがあるが、話者識別と話者照合において固有空間を作成する手順は本質的に同じである。したがって、図2 および図3 において同一または相当部分には同じ参照符号を付している。

【００４０】ステップ２４において、学習用データ２２に表された話者の各々について学習用話者モデルを発展させ教え込む。その結果、各話者モデルの集合２６が生成される。ここでは、隠れマルコフモデルについて示したが、これに限定されるものではなく、連接に適したパラメータを有する音声モデルであればどのようなモデルでもよい。好ましくは、モデルにより画定された全ての音声単位が少なくとも一度は各話者の実際の音声によって教え込まれるよう十分な学習用データを使ってモデル２６を学習させる。図2 および図3 には明確に示していないが、モデルを洗練するのに適した話者適応手順２４を付加的に含めることができる。このような付加的な手順の例としては、最大事後推定法(Ｍａｘｉｍｕｍ Ａｐｏｓｔｔｔｔｅｒｉｏｒｉ ｅｓｔｉｍａｔｉｏｎ：ＭＡＰ) や、最大線形回帰法(ＭＬＬＲ) などの変換に基づく手法が挙げられる。

【００４１】話者モデル２６を作成する目的は、学習用データ集合を正確に表し、各学習用話者を配置し新たな話者の発声を検査する固有空間の境界をこの集合を使用して画定することにある。

【００４２】モデル２６を作成した後、ステップ２８において、各話者についてのモデルを使用してスーパーベクトル３０を作成する。スーパーベクトル３０は、各話者についてのモデルのパラメータを連接させて構成することができる。隠れマルコフモデルを使用する場合、各話者についてのスーパーベクトルは、パラメータ (一般に浮動小数点数) の配列リストとなる。これらのパラメータは、その話者についての隠れマルコフモデルのパラメータの少なくとも一部に対応する。与えられた話者についてのスーパーベクトルには、各音声単位に対応するパラメータが含まれる。パラメータは都合のよい順序に編成することができる。その順序は重要ではないが、一旦ある順序が採用されると学習用話者全員についてその順序に従わせる必要がある。

【００４３】スーパーベクトルを作成するために使用するモデルパラメータの選択は、利用できるコンピュータシステムの処理能力に依存する。隠れマルコフモデルを使用した場合、ガウス平均値(ｔｈｅ Ｇａｕｓｓｉａｎ ｍｅａｎｓ) からスーパーベクトルを作成することにより良い結果が得られた。もし、さらに大きな処理能力を利用できるならば、スーパーベクトルに他のパラメータ (例えば、図1 に示す遷移確率１２、パラメータ１８中の共分散行列Uj など) を含めることができる。もし、隠れマルコフモデルにより離散的な出力(確率密度と対照的な) が生成されるならば、これらの出力値を使用してスーパーベクトルを作成することができる。

【００４４】スーパーベクトルを作成した後、ステップ３２において、次元数削減演算を行う。次元数削減は、元の高次元のスーパーベクトルを基底ベクトルに変える

どのような線形変換を通じても達成できる。不完全ではあるが例を挙げると、主成分分析（Ｐｒｉｎｃｉｐａｌ Ｃｏｍｐｏｎｅｎｔ Ａｎａｌｙｓｉｓ：ＰＣＡ）、独立成分分析（Ｉｎｄｅｐｅｎｄｅｎｔ Ｃｏｍｐｏｎｅｎｔ Ａｎａｌｙｓｉｓ：ＩＣＡ）、線形識別分析（Ｌｉｎｅａｒ Ｄｉｓｃｒｉｍｉｎａｔｅ Ａｎａｌｙｓｉｓ：ＬＤＡ）、因子分析（Ｆａｃｔｏｒ Ａｎａｌｙｓｉｓ：ＦＡ）、特異値分析（Ｓｉｎｇｕｌａｒ Ｖａｌｕｅ Ｄｅｃｏｍｐｏｓｉｔｉｏｎ：ＳＶＤ）などが挙げられる。

【００４５】特に、本発明を実行する際に有用な次元数削減手法を以下に示す。音声認識に関する話者独立型モデルから得られたＴ個の学習用スーパーベクトルにより構成される一つの集合を考える。これらのスーパーベクトルの各々は次元数Ｖを有すると仮定する。従って、全てのスーパーベクトルをＸ＝[ｘ１，ｘ２，…，ｘＶ]＾Ｔ（Ｖ＊1 ベクトル）のように表すことができる。次元数Ｅの新たなベクトルを生成するために、スーパーベクトルに適用可能な線型変換Ｍを考える。ここで、Ｅ≦Ｔである。Ｔは、学習用スーパーベクトルの数である。変換されたベクトルの各々は、Ｗ＝[ｗ１，ｗ２，…，ｗＥ]＾Ｔのように表すことができる。線型変換Ｍのパラメータの値は、Ｔ個の学習用スーパーベクトルによる集合から何らかの方法で計算される。

【００４６】このようにして、線型変換Ｗ＝Ｍ＊Ｘが得られる。ＭはＥ＊Ｖの次元数を有し、ＷはＥ＊1 の次元数を有する。ここで、Ｅ≦Ｔである。Ｔ個のスーパーベクトルによる集合のうち特別のものについては、Ｍは定数になる。 Ｗは次元数Ｅ（Ｅ≦Ｔである。）を有するため、Ｔ個のスーパーベクトルによる１つの集合から線形変換Ｍを計算するためにいくつかの次元数削減手法を使用することができる。例として、主成分分析（Ｐｒｉｎｃｉｐａｌ Ｃｏｍｐｏｎｅｎｔ Ａｎａｌｙｓｉｓ）、独立成分分析（Ｉｎｄｅｐｅｎｄｅｎｔ Ｃｏｍｐｏｎｅｎｔ Ａｎａｌｙｓｉｓ）、線形識別分析（Ｌｉｎｅａｒ Ｄｉｓｃｒｉｍｉｎａｎｔ Ａｎａｌｙｓｉｓ）、因子分析（Ｆａｃｔｏｒ Ａｎａｌｙｓｉｓ）、特異値分析（Ｓｉｎｇｕｌａｒ Ｖａｌｕｅ Ｄｅｃｏｍｐｏｓｉｔｉｏｎ）がある。

【００４７】この発明は、例に挙げた方法に限らず、入力ベクトルが話者依存型モデルにより得られた学習用スーパーベクトルであるという特別のケースにおいて不変線型変換Ｍを見つけるためのどのような方法を使用しても行うことができる。ここでＭは前記手法を行うために使用される。

【００４８】ステップ３２において生成された基底ベクトルは、固有ベクトルにより張られる固有空間を定める。次元数削減により、学習用話者一人当たり一つの固有ベクトルが作成される。したがって、Ｔ人の学習用話者が存在するときは、次元数削減ステップ３２によりＴ個の固有ベクトルが生成される。これらの固有ベクトルにより、この説明において固有音声空間あるいは固有空間と呼ぶ空間が定められる。

【００４９】固有空間を構成する固有ベクトルの各々は、図2 および図3 の３４に示すように、それぞれ異なった次元を表し、それに沿って異なる話者を区別することができる。元の学習用集合の中の各スーパーベクトルは、これら固有ベクトルの線形結合として表すことができる。固有ベクトルは、データをモデル化する際の重要性に応じて配列される。第一の固有ベクトルは第二の固有ベクトルよりも重要であり、第二の固有ベクトルは第三の固有ベクトルよりも重要である、という具合である。実験によれば、第一の固有ベクトルは男女を表す次元に対応する。

【００５０】ステップ３２においては最大Ｔ個の固有ベクトルを作成するが、実際にはこれらの固有ベクトルのいくつかを捨てて最初のＮ個の固有ベクトルだけを保持することもできる。ステップ３６においては、Ｔ個の固有ベクトルのうちＮ個を選択的に抽出してパラメータ数を削減した固有空間３８を作成する。より高位に配列された固有ベクトル（前述の第一の固有ベクトルに対する第二、第三の固有ベクトル）は、一般に話者間の識別のための重要な情報を比較的含んでいないため捨てることができる。固有音声空間を縮小して学習用話者の総数よりも小さくすることにより、限られた記憶容量と処理装置による実用的なシステムを構築する際に有用な固有のデータ圧縮が行える。

【００５１】学習用データから固有ベクトルを作成した後、学習用データにおける各話者を固有空間内に表す。話者識別を行う場合には、図2 に示すステップ４０ａにおいて、各登録話者を固有空間内に表す。これを４２ａに図式的に示す。話者照合を行う場合には、図3 に示すステップ４０ｂにおいて、登録話者および仮想の詐称者を固有空間内に表す。これを４２ｂに図式的に示す。話者は、図2 の４２ａに示すように固有空間内に点として、あるいは図3 の４２ｂに示すように固有空間内に確率分布として表す。

【００５２】＜話者識別あるいは話者照合システムの使用＞図4 は、この発明の実施の形態による話者識別システムおよび話者照合システムの使用を説明するためのフローチャートである。図4 を参照して、ステップ４４において、話者識別あるいは話者照合を求めるユーザは新たな音声データを提供する。ステップ４６において、新たなデータを使用して話者依存型モデル４８を学習させる。ステップ５０において、モデル４８を使用してスーパーベクトル５２を作成する。なお、新たな音声データは各音声単位の例を必ずしも含んでいない。例えば、新たな発話が非常に短いためにすべての音声単位の例を含んでいないかもしれない。本システムは、この問題を解決する。

11

【００５３】ステップ５４において、スーパーベクトル５２に対して次元数削減を行う。その結果ステップ５６において、固有空間内に新たなデータが図4 中の５８に示すように位置付けられる。図4 中５８では、学習用データに基づいて固有空間内に既存データを示す部分はドットで表し、新たな音声データはスターマークで表している。

【００５４】固有空間内に新たなデータを位置付けた後、学習用話者に対応して既存データを示す点または分布との類似度の評価が行われる。図4 には、話者識別および話者照合の典型的な実施例を示している。

【００５５】話者識別の場合には、ステップ６２において、新たな音声データは固有空間内の最も近い学習用話者に割り当てられる。この様子を図4 中の６４に示す。

【００５６】システムはこのようにして、新しい音声と固有空間内におけるデータ点またはデータ分布が最も近い既存の学習用話者とが同一であると認識する。

【００５７】話者照合の場合には、ステップ６６において、システムは新しいデータを示す点を調べて、それが固有空間内の登録話者に対してあらかじめ定められたしきい値類似度の範囲内にあるかどうかを決定する。ステップ６８において、新たな話者データが登録話者よりも詐称者に近いときは、予防手段としてシステムはそのデータを棄却する。この様子を図4 中の６９に示す。ここには登録話者に対する類似度および詐称者に対する類似度が描かれている。

【００５８】＜最尤固有空間分析手法（Ｍａｘｉｍｕｍ Ｌｉｋｅｌｉｈｏｏｄ Ｅｉｇｅｎｓｐａｃｅ Ｄｅｃｏｍｐｏｓｉｔｉｏｎ Ｔｅｃｈｎｉｑｕｅ：ＭＬＥＤ）＞新たな話者を固有空間内に位置付ける一つの簡単な手法は、単純射影演算を用いることである。射影演算により、固有空間外の点にできるだけ近い固有空間内の点が新たな話者による入力音声に対応する。これらの点が実際にはHMMの集合を再構成することができるスーパーベクトルであることは重要ではない。

【００５９】射影演算は比較的未熟な手法であるので、固有空間内の地点が新たな話者に関して最適であるということは保証されない。さらに、射影演算においては、新たな話者についての完全なHMM集合を表すための完全なデータ集合が新たな話者についてのスーパーベクトルに含まれていることが必要とされる。この要求により、実用上の制限をかなり受ける。射影を使用して新たな話者を固有空間内に位置付ける場合、新たな話者は、全ての音声単位がデータ内に表されるように十分な入力音声を提供しなければならない。例えば、隠れマルコフモデルにより英語のすべての音素を表そうとすると、単純射影手法を使用する前に学習用話者は全ての音素の例を提供しなければならない。応用する際にこのような制限が存在することは実用的でない。

【００６０】この実施の形態における最尤手法は、上述

12

の単純射影の欠点の両方に焦点をあてたものである。最尤手法では、新たな話者によって供給される音声の最大生成確率を有する隠れマルコフモデルに対応したスーパーベクトルを表す点を固有空間内に見つける。

【００６１】単純射影演算ではスーパーベクトルの全ての要素を同等の重要性を有するものとして取り扱うのに対して、最尤手法では、実際の適用データから生じる確率に基づいてより起こりそうなデータにはより大きな重みをつけるようにする。単純射影演算と違って、たとえ新たな話者により十分な学習用データ集合が提供されない場合であっても最尤手法は機能する。すなわち、音声単位のいくつかのデータが欠けている場合である。実際には、最尤手法ではスーパーベクトルが作成された状況を考慮に入れる。すなわち、他のモデルよりも新たな話者が提供した入力音声を生成しやすいという確率を有する隠れマルコフモデルからスーパーベクトルを作成する。実用上は、入力音声が実際にどのくらい利用できるかの程度にかかわらず、最尤手法は、固有空間内において新たな話者の入力音声に最も一致するスーパーベクトルを選択する。ここで説明の便宜上、新たな話者はアラバマ出身の若い女性であると仮定する。最尤手法では、この話者から発せられた数音節に基づいて、アラバマ出身女性のアクセントに一致する全ての音素を表す部分が固有空間内において選択される。

【００６２】図5 は、最尤手法がどのように行われるかを説明するための図である。図5 を参照して、新たな話者からの入力音声を使用してスーパーベクトル７０を作成する。上述のように、スーパーベクトルは、ケプストラム係数または同様のものに対応した音声パラメータの連接リストを含む。これらのパラメータは、新たな話者に対応した隠れマルコフモデルから抽出されたガウス平均を表す浮動小数点数である。他の隠れマルコフモデル平均を使用することもできる。これらの隠れマルコフモデル平均は、図5 中の７２のドットで表される。データが全て揃っている場合、スーパーベクトル７０は、隠れマルコフモデル平均の各々についての浮動小数点数を含み、これらは隠れマルコフモデルによって表された音声単位の各々に対応している。ここで、音素”ａｈ”についてのパラメータは存在するが音素”ｉｙ”についてのパラメータが欠けている場合を仮定する。

【００６３】固有空間３８は、固有ベクトル７４，７６，７８の集合によって表される。固有ベクトルの各々を、図5 中にW1，W2，・・・，Wnとして示された対応する固有値と掛け合わせることにより、新たな話者からの観測データに対応したスーパーベクトル７０を固有空間内に表すことができる。最初これらの固有値は未知である。最尤手法はこれら未知の固有値の値を見つける。さらにいうと、これらの値は、固有空間内で新たな話者を最もよく表す最適解を探すことにより選ばれる。固有値を固有空間３８内の対応する固有ベクトルと掛け

13

合わせた後、それらの結果を足しあわせて適応モデルを表すスーパーベクトル８０を作成する。入力話者のスーパーベクトル７０はいくつかのパラメータ値（例えば、”ｉｙ”パラメータ）を欠いているのに対して、適応モデルを表すスーパーベクトル８０では全ての値が揃っている。これはこの発明により得られる一つの利益である。さらに、スーパーベクトル８０内の値は最適解、すなわち固有空間内に新たな話者を表す最大尤度を表す。

【００６４】各固有値W1，W2，・・・，Wnは、最尤ベクトルを含んでいるとみなすことができる。図5中の８２に最尤ベクトルの概略を示す。図5に示すように、最尤ベクトル８２は最尤ベクトル８２は固有値W1，W2，・・・，Wnの集合を含む。

【００６５】図6は、最尤手法を使用した適応化の手順を説明するためのフローチャートである。図6を参照して、まずステップ１００において、観測データを含む新たな話者からの音声を使用して隠れマルコフモデル集合１０２を作成する。ステップ１０４において、隠れマルコフモデル集合１０２を使用してスーパーベクトル１０６を作成する。スーパーベクトル１０６は、隠れマルコフモデル１０２から抽出された隠れマルコフモデルパラメータの連接リストを含む。

【００６６】ステップ１０８において、スーパーベクトル１０６を使用して確率関数Qを作成する。この実施の形態では、確率関数は、あらかじめ定義された隠れマルコフモデル１０２に関して観測されたデータの生成確率を表す。

【００６７】確率関数Qが確率項Pだけでなくその対数項ｌｏｇＰも含んでいるならば、確率関数Qの後の扱いが容易になる。

【００６８】ステップ１１０において、固有値W1，W2，・・・，Wnの各々について個々に確率関数Qの導関数を求めることにより確率関数Qの最大化を行う。例＊

$$Q(\lambda,\hat{\lambda}) = \text{const} - \frac{1}{2}P(O|\lambda)\sum_{\substack{\text{states}\\ \text{in}\,\lambda}}^{S\lambda}\sum_{\substack{\text{mixt}\\ \text{gauss}\\ \text{in}\,S}}^{Ms}\sum_{\substack{\text{time}\\ t}}^{T}\left\{\gamma_m^{(s)}(t)[n\log(2\pi)+\log|C_m^{(s)}|+h(o_t,m,s)]\right\}$$

ここで、

【００７３】

【数3】

$$h(o_t,m,s) = (o_t-\hat{\mu}_m^{(s)})^T C_m^{(s)-1}(o_t-\hat{\mu}_m^{(s)})$$

であり、さらに、

【００７４】

【数4】

14

＊えば、固有空間の次元数が１００の場合、このシステムでは、確率関数Qの１００個の導関数を求めてそれぞれを0と置いてそれぞれのWを求める。これは計算量が多いように見えるが、何千もの計算を行うことが一般的に要求される従来のMAP法あるいはMLLR法に比べてはるかに計算量が少ない。

【００６９】このようにして得られたWの集合は、最尤点に対応した固有空間内の点の認定に必要とされる固有値を表す。したがって、Wの集合は固有空間内の最尤ベクトルを含む。固有ベクトルの各々（図5に示す固有ベクトル７４，７６，７８）は、直交ベクトル集合または直交座標集合を定義し、この集合に対して固有値が掛け合わされて固有空間内に制限された点が定義される。ステップ１１２において、この最尤ベクトルを使用して固有空間内の最適点（図4中の６６に示す）に対応したスーパーベクトル１１４が作成される。ステップ１１６において、スーパーベクトル１１４を使用して新たな話者についての適応モデル１１８を作成する。

【００７０】この実施の形態における最尤法において、モデルλに関する観測値O（$O=o_1,\cdots,o_T$）の尤度を最大にすることが望まれる。これは、数1に示す補助関数Qの最大化を反復することにより行うことができる。数1において、λは反復処理時の現行モデルを表し、$\hat{\lambda}$は推定モデルを表す。

【００７１】

【数1】

$$Q(\lambda,\hat{\lambda}) = \sum_{\theta\in states} P(O,\theta|\lambda)\log[P(O,\theta|\hat{\lambda})]$$

予備の近似計算として、平均値のみについて最大化を実行してもよい。隠れマルコフモデル集合により確率Pが与えられた場合には、以下の数2を得る。

【００７２】

【数2】

$$Q(\lambda,\hat{\lambda}) = \text{const} - \frac{1}{2}P(O|\lambda)\sum_{\substack{\text{states}\\ \text{in}\,\lambda}}^{S\lambda}\sum_{\substack{\text{mixt}\\ \text{gauss}\\ \text{in}\,S}}^{Ms}\sum_{\substack{\text{time}\\ t}}^{T}\left\{\gamma_m^{(s)}(t)[n\log(2\pi)+\log|C_m^{(s)}|+h(o_t,m,s)]\right\}$$

o_t　：時刻tにおける特徴ﾍﾞｸﾄﾙ

$C_m^{(s)-1}$　：状態Sの混合ｶﾞｳｼｱﾝmについての反共分散

$\hat{\mu}_m^{(s)}$　：状態S，混合成分mの近似適応化平均

$\gamma_m^{(s)}(t)$　：P（混合ｶﾞｳｼｱﾝm|λ，o_tを使用）

である。

【００７５】新たな話者の隠れマルコフモデルについてのガウス平均が固有空間内に配置されていると仮定する。数5に示す平均スーパーベクトル／μ_j（j＝1，・・・，E）によってこの空間を広げる。

15

【００７６】
【数5】

$$\overline{\mu_j} = \begin{bmatrix} \overline{\mu}_1^{(1)}(j) \\ \overline{\mu}_2^{(1)}(j) \\ \cdot \\ \cdot \\ \overline{\mu}_m^{(s)}(j) \\ \overline{\mu}_{Ms_\lambda}^{(S_\lambda)}(j) \end{bmatrix}$$

ここで、$\overline{\mu}_m^{(s)}(j)$は、固有ベクトル（固有モデル）jの状態sにおける混合ガウシアンmについての平均ベクトルを表す。
【００７７】したがって、以下の数6に示される$\hat{\mu}$を必要とする。
【００７８】
【数6】

$$\hat{\mu} = \sum_{j=1}^{B} w_j \overline{\mu}_j$$

$$\frac{\partial Q}{\partial w_e} = 0 = \sum_{\substack{\text{states} \\ \text{in } \lambda}}^{S_\lambda} \sum_{\substack{\text{mixt} \\ \text{gauss} \\ \text{in } S}}^{MS} \sum_{\substack{\text{time} \\ t}}^{T} \left\{ \frac{\partial}{\partial w_e} \gamma_m^{(s)}(t) h(o_t, s) \right\}, e = 1..E.$$

上記導関数を計算することにより、数10を得る。
【００８３】

$$0 = \sum_s \sum_m \sum_t \gamma_m^{(s)}(t) \left\{ -\overline{\mu}_m^{(s)T}(e) C_m^{(s)-1} o_t + \sum_{j=1}^{E} w_j \overline{\mu}_m^{(s)T}(j) C_m^{(s)-1} \overline{\mu}_m^{(s)}(e) \right\}$$

これにより、数11に示される一群の線形方程式を得る。

$$\sum_s \sum_m \sum_t \gamma_m^{(s)}(t) \overline{\mu}_m^{(s)T}(e) C_m^{(s)-1} o_t = \sum_s \sum_m \sum_t \gamma_m^{(s)}(t) \sum_{j=1}^{E} w_j \overline{\mu}_m^{(s)T}(j) C_m^{(s)-1} \overline{\mu}_m^{(s)}(e), e = 1..E.$$

＜固有空間内における類似度評価＞話者を固有空間内に点として表す場合は、単純な幾何学的距離の計算を用いて、新たな話者に対してどの学習用話者が一番近いかを認定することができる。話者を固有空間内に分布として表す場合は、新たな話者データを観測値Oとして取り扱い、各分布候補（学習用話者を表す）を検査することによって類似度を評価してその候補が観測データを生成した確率を決定する。最も高い確率を有する候補が最も近い類似度を有するものと評価する。高い安全性を必要とする応用例においては、最も高い確率を有する候補が、あらかじめ定められたしきい値よりも低い確率値を有する場合には、照合を棄却することが望まれる。費用関数を用いて高度の確実性を欠く候補を除外することができる。
【００８５】上述のように、新たな話者の学習用話者に対する類似度の評価は、完全に固有空間内において行われる。さらに高度の正確さを得るために、これに代えてベイズ推定法を用いることができる。

16

＊／μ_jは直交しており、w_jは話者モデルの固有値である。どんな新たな話者であっても、観測された話者のデータベースの線形結合によりモデル化することができると仮定する。
【００７９】
【数7】

$$\hat{\mu}_m^{(s)} = \sum_{j=1}^{E} w_j \overline{\mu}_m^{(s)}(j)$$

Qを最大化するため、以下の処理を行う。
【００８０】
【数8】

$$\frac{\partial Q}{\partial w_e} = 0, \quad e = 1..E.$$

なお、固有ベクトルは直交するため、（$\partial w_i / \partial w_j$）$=0$，$i \neq j$である。
＊【００８１】したがって、数9を得る。
【００８２】
【数9】

$$\frac{\partial Q}{\partial w_e} = 0 = \sum_{\substack{\text{states} \\ \text{in } \lambda}}^{S_\lambda} \sum_{\substack{\text{mixt} \\ \text{gauss} \\ \text{in } S}}^{MS} \sum_{\substack{\text{time} \\ t}}^{T} \left\{ \frac{\partial}{\partial w_e} \gamma_m^{(s)}(t) h(o_t, s) \right\}, e = 1...E.$$

※【数10】

$$0 = \sum_s \sum_m \sum_t \gamma_m^{(s)}(t) \left\{ -\overline{\mu}_m^{(s)T}(e) C_m^{(s)-1} o_t + \sum_{j=1}^{E} w_j \overline{\mu}_m^{(s)T}(j) C_m^{(s)-1} \overline{\mu}_m^{(s)}(e) \right\}$$

★【００８４】
【数11】

$$\sum_s \sum_m \sum_t \gamma_m^{(s)}(t) \overline{\mu}_m^{(s)T}(e) C_m^{(s)-1} o_t = \sum_s \sum_m \sum_t \gamma_m^{(s)}(t) \sum_{j=1}^{E} w_j \overline{\mu}_m^{(s)T}(j) C_m^{(s)-1} \overline{\mu}_m^{(s)}(e), e = 1..E.$$

【００８６】ベイズ推定法を用いた類似度評価を高めるために、固有空間内の学習用話者のガウス密度に対して、次元数削減を通じて捨てられた話者データを表す直交補空間内の推定限界密度を掛け合わせる。話者モデルであるスーパーベクトルに基づいて次元数削減を実行すれば高次元空間から低次元空間にデータを大幅に圧縮できることがこれによりわかるであろう。次元数削減によって最も重要な基底ベクトルは保持されるが、より上位の情報のいくつかは捨てられる。ベイズ推定法は、この捨てられた情報に対応した限界ガウス密度を推定する。元の固有空間は、スーパーベクトルの次元数削減処理を通じての線形変換により作成される。ここではN個の全成分からM個の成分が抽出される。抽出されるM個の成分が少ないほど、最大限の固有値に対応した変換基底の

17

より低次元の下位空間を表すことができる。このようにして、重要でない成分i（ｉ＝M+1，・・・，N）は捨てられるのに対し、成分i（ｉ＝1，・・・，M）によって固有空間が定義される。これら二つの成分集合は、相互に排他的で補完的な二つの下位空間を定義する。主要な下位空間は重要な固有空間を表し、その直交成分は次元数削減を通じて捨てられたデータを表す。

【００８７】これら二つの各直交空間内のガウス密度の積として、数12に示す式により尤度推定値を計算することができる。

【００８８】

【数12】

$$\hat{P}(x|\Omega) = P_E(x|\Omega)^* P_{\bar{E}}(x|\Omega)$$

数12において、第1項は固有空間E内の単ガウス密度であり、第2項は固有空間に対して直交する空間内の単ガウス分布である。固有空間への射影と残差だけを使用して二つの項を完全に学習用データベクトル集合から推定できることがわかる。

【００８９】

【発明の効果】この発明に従った音声評価方法は、固有空間内において新たな音声データと学習用データとの類似度を評価するため以下の利点がある。

【００９０】第一に、固有空間は、単に選択された数個の特徴だけでなく、各話者の全体を簡潔、低次元の方法で表す。

【００９１】また、固有空間内に含まれる次元数は、元の話者モデル空間あるいは特徴ベクトル空間内に含まれるよりも一般にかなり少ないので、固有空間内で実行される類似度の計算を非常に速くすることができる。

【００９２】また、システムにおいては、元の学習用データを作成するのに使用したすべての例、発話が新たな

18

音声データに含まれていることは必要とされない。この発明によれば、その構成要素の一部を欠くスーパーベクトルに対して次元数の削減を行うことができる。その結果としての固有空間内の分布点は話者を明確に表す。

【図面の簡単な説明】

【図1】隠れマルコフモデルの典型例を説明するための図である。

【図2】この発明の実施の形態による話者識別システムを実行するための固有空間の作成を説明するためのフローチャートである。

【図3】この発明の実施の形態による話者照合システムを実行するための固有空間の作成を説明するためのフローチャートである。

【図4】この発明の実施の形態による話者識別システムおよび話者照合システムの使用を説明するためのフローチャートである。

【図5】最尤法がどのように行われるかを説明するための図である。

【図6】最尤法を使用した適応化の手順を説明するためのフローチャートである。

【符号の説明】

２６　各話者モデルの集合

３０，５２，７０，１０６，１１４　スーパーベクトル

４２ａ　登録話者

４２ｂ　登録話者および仮想の詐称者

４８　話者依存型モデル

７４，７６，７８　固有ベクトル

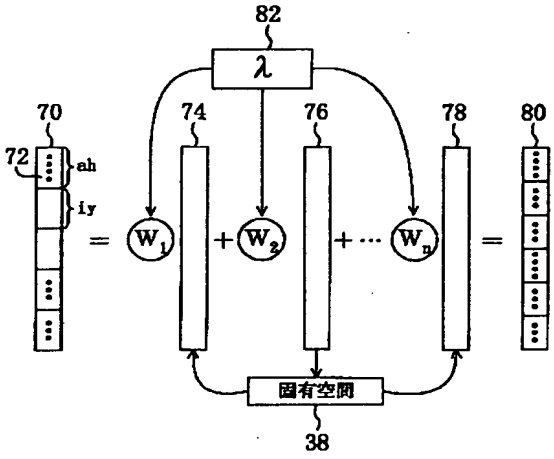
８０　適応モデルを表すスーパーベクトル

８２　最尤ベクトル

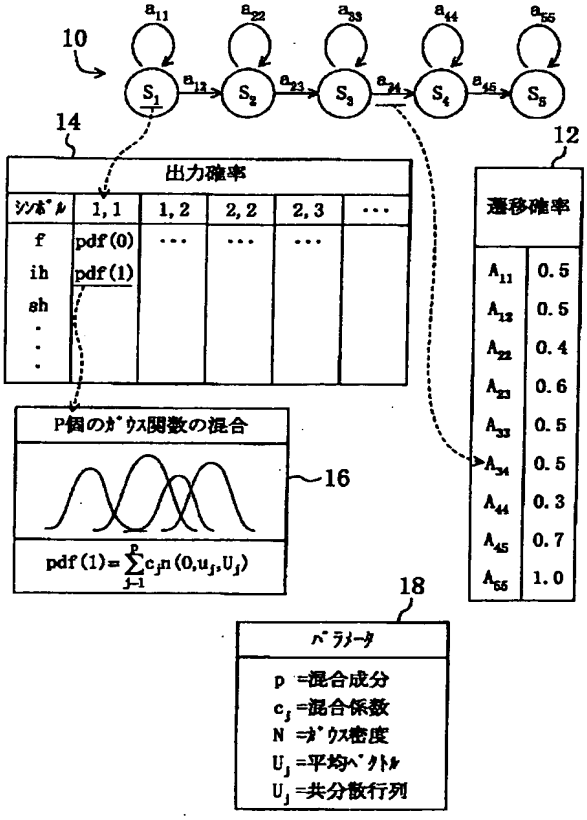
１０２　隠れマルコフモデル集合

１１８　新たな話者についての適応モデル

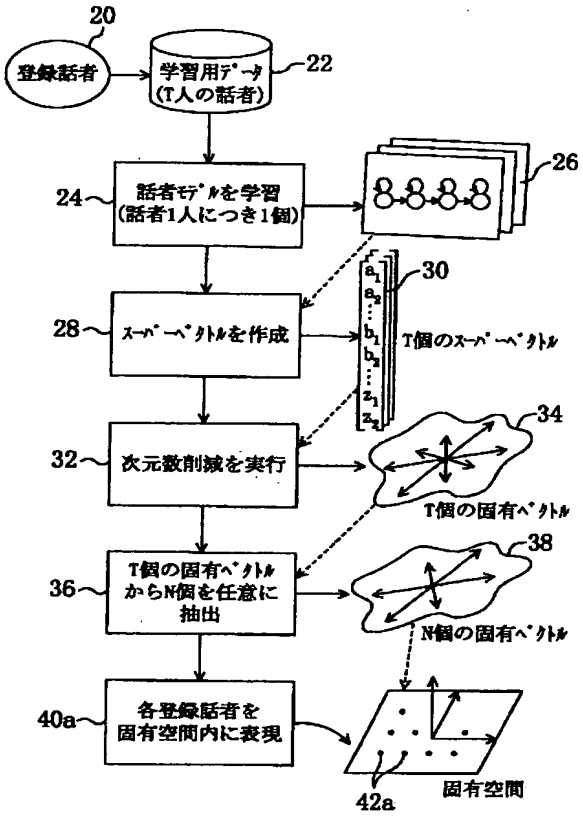
【図5】



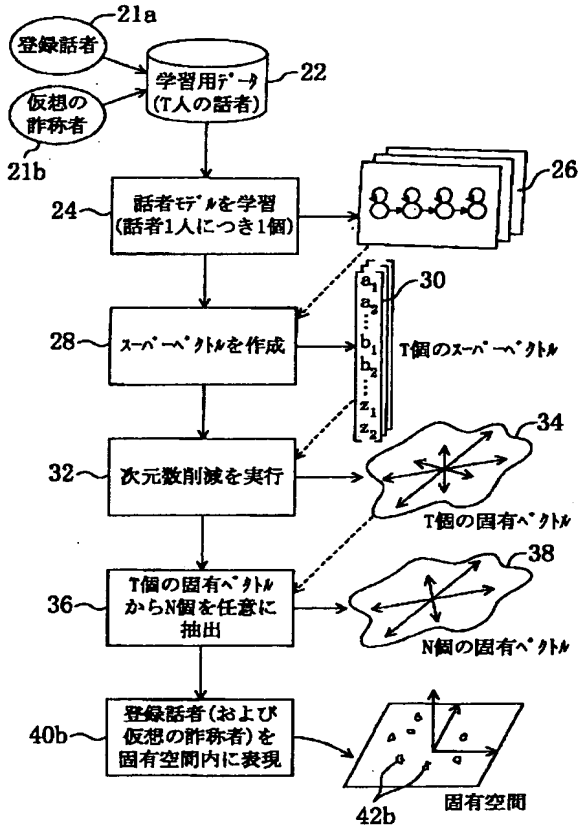
【 図1 】


10
14
12
16
18
出力確率
遷移確率
P個のｶﾞｳｽ関数の混合
ﾊﾟﾗﾒｰﾀ
p =混合成分
c_j =混合係数
N =ｶﾞｳｽ密度
U_j =平均ﾍﾞｸﾄﾙ
U_j =共分散行列

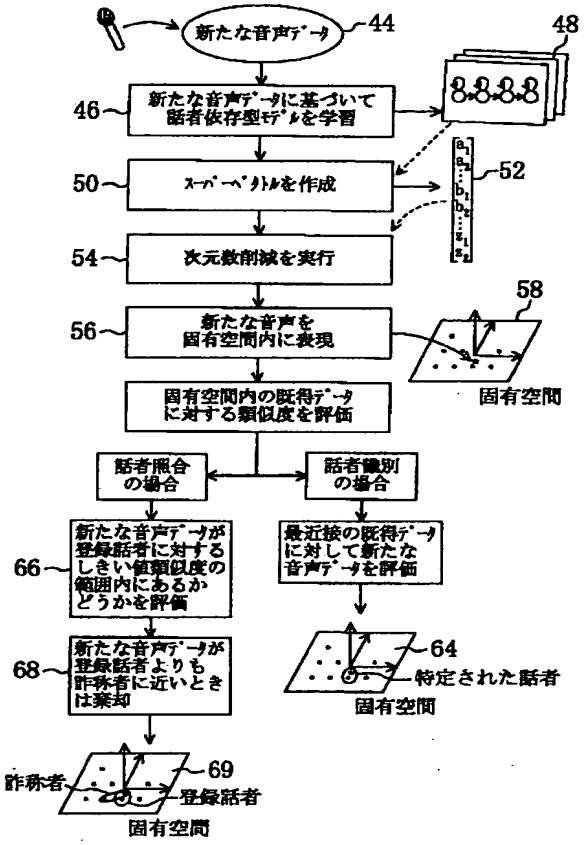
【 図2 】


20
登録話者
学習用ﾃﾞｰﾀ
(T人の話者)
22
24
話者ﾓﾃﾞﾙを学習
(話者1人につき1個)
26
28
ｽｰﾊﾟｰﾍﾞｸﾄﾙを作成
30
T個のｽｰﾊﾟｰﾍﾞｸﾄﾙ
32
次元数削減を実行
34
T個の固有ﾍﾞｸﾄﾙ
36
T個の固有ﾍﾞｸﾄﾙ
からN個を任意に
抽出
38
N個の固有ﾍﾞｸﾄﾙ
40a
各登録話者を
固有空間内に表現
42a
固有空間

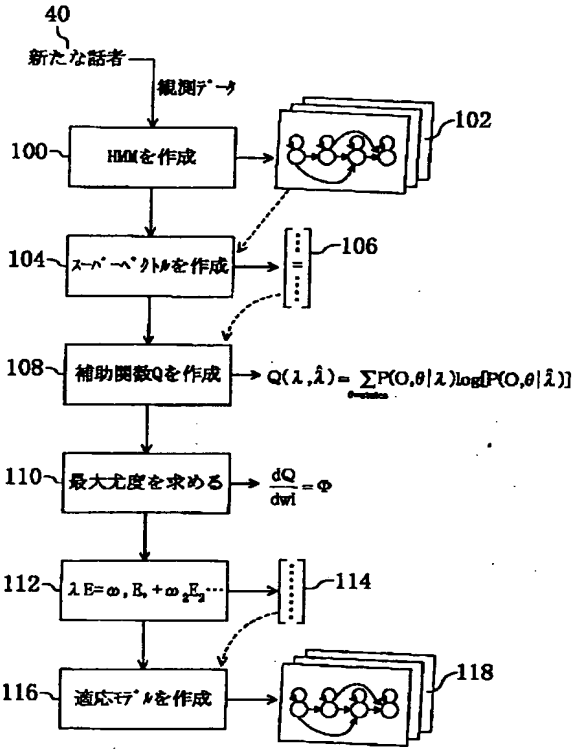
【 図3 】


21a
登録話者
仮想の詐称者
21b
学習用ﾃﾞｰﾀ(T人の話者)
22
話者ﾓﾃﾞﾙを学習(話者1人につき1個)
24
26
28
ｽｰﾊﾟｰﾍﾞｸﾄﾙを作成
30
T個のｽｰﾊﾟｰﾍﾞｸﾄﾙ
32
次元数削減を実行
34
T個の固有ﾍﾞｸﾄﾙ
36
T個の固有ﾍﾞｸﾄﾙからN個を任意に抽出
38
N個の固有ﾍﾞｸﾄﾙ
40b
登録話者(および仮想の詐称者)を固有空間内に表現
42b
固有空間

【 図4 】


新たな音声ﾃﾞｰﾀ
44
48
46
新たな音声ﾃﾞｰﾀに基づいて話者依存型ﾓﾃﾞﾙを学習
50
ｽｰﾊﾟｰﾍﾞｸﾄﾙを作成
52
54
次元数削減を実行
56
新たな音声を固有空間内に表現
58
固有空間
固有空間内の既得ﾃﾞｰﾀに対する類似度を評価
話者照合の場合
話者識別の場合
66
新たな音声ﾃﾞｰﾀが登録話者に対するしきい値類似度の範囲内にあるかどうかを評価
最近接の既得ﾃﾞｰﾀに対して新たな音声ﾃﾞｰﾀを評価
64
特定された話者
固有空間
68
新たな音声ﾃﾞｰﾀが登録話者よりも詐称者に近いときは棄却
詐称者
69
登録話者
固有空間

【 図6 】



フロントページの続き

(72)発明者　ジュンクア　ジーン－クランデ
アメリカ合衆国　カリフォルニア州
93111　サンタ　バーバラ，サンタ　アナ
アヴェニュー　146